120

AF $

Attorney's Docket No.: 13865-023001

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant : Jonathan H. Young et al. Art Unit : 2654
Serial No. : 09/390,370 Examiner : Angela A. Armstrong
Filed : September 7, 1999
Title : EXPANDING AN EFFECTIVE VOCABULARY OF A SPEECH RECOGNITION SYSTEM

Mail Stop Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

APPEAL BRIEF

(1) Real Party in Interest

ScanSoft, Inc., the assignee of this application, is the real party in interest.

(2) Related Appeals and Interferences

There are no related appeals or interferences.

(3) Status of Claims

Claims 1-9 and 11-55 are pending in this application, with claims 1, 34, 35, 37, 46, and 51 being independent. Appellant has appealed the rejection of claims 1-9 and 11-55.

(4) Status of Amendments

The claims were not amended subsequent to the final rejection.

(5) Summary of claimed subject matter

The following summarizes each independent claim with references to the application specification and drawings. The references to the specification and drawings are meant to be exemplary, and not limiting.

Independent claim 1 is directed to a method of expanding an effective active vocabulary of a speech recognition system. As shown in Fig. 29, one or more recognition candidates are received (step 2900) from a speech recognizer that performs speech recognition using a set of

acoustic models representative of an active vocabulary of the system. See the specification at page 68, lines 14-16 and Figs. 29 and 30.

When the received recognition candidate includes a word fragment (step 2905), a determination is made as to whether the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system using a spelling rule associated with the word fragment (steps 2915, 3010, 3015). See the specification at page 68, line 20 to page 70, line 15 and Figs. 29 and 30. As a result of using the associated spelling rule, a spelling of the proposed word differs from a spelling that would result from merely concatenating the particular word fragment with the one or more adjacent word fragments or words. See the specification at page 4, lines 27-31; page 51, line 29 to page 53, line 9; and page 75, lines 4-19.

If the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system (step 3025), the recognition candidate is modified to substitute the proposed word for the word fragment and the one or more adjacent word fragments or words used to form the proposed word (step 3030). See the specification at page 70, lines 22-29 and Figs. 29 and 30. If the word fragment may not be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system (steps 3025, 3045), the recognition candidate is discarded. See the specification at Figs. 29 and 30.

Independent claim 34 is directed to a method of recognizing speech. A set of one or more recognition candidates are received (step 2900) from a speech recognizer that uses a set of acoustic models representative of an active vocabulary. See the specification at page 68, lines 14-16 and Figs. 29 and 30. The set of acoustic models includes models of words, models of roots that are not words, and models of affixes that are not words. The affixes include prefixes and suffixes. See the specification at page 41, line 31 to page 42, line 20; page 47, lines 20-25; and Figs. 14A-15C.

When a received recognition candidate includes an affix (step 2905), the affix is combined with one or more adjacent words, roots, or other affixes to form a new word (steps 2915, 3010, 3015) and the recognition candidate is modified to substitute the new word for the affix and the one or more adjacent words, roots, or other affixes used to form the new word (step

3030). See the specification at page 68, line 20 to page 70, line 15 and lines 22-29 and Figs. 29 and 30. Formation of the new word includes using a spelling rule associated with the affix that causes the spelling of the new word to differ from a spelling that would result from merely concatenating the affix with the one or more adjacent words, roots, or other affixes. See the specification at page 4, lines 27-31; page 51, line 29 to page 53, line 9 and page 75, lines 4-19.

Independent claim 35 is directed to a method of generating an acoustic model of a word fragment. See the specification at page 2, lines 22-26 and page 3, lines 20-26. A word of an active vocabulary is compared to a similar word of a backup dictionary to identify a word fragment that may be used to convert the word of the active vocabulary to the word of the backup dictionary. See the specification at page 54, line 16 to page 55, line 4. The acoustic model of the word fragment is generated using a portion of an acoustic model of the word of the backup dictionary that is not included in an acoustic model of the word of the active vocabulary. See the specification at page 1, line 33 to page 2, line 10; page 10, lines 7-16; page 20, lines 12-16; and page 51, lines 3-7.

Independent claim 37 is directed to a method of generating acoustic models of word fragments. See the specification at page 2, lines 22-26 and page 3, lines 20-26. Words of an active vocabulary are compared to similar words of a backup dictionary to identify spelling rules that may be used to convert the words of the active vocabulary to words of the backup dictionary. See the specification at page 53, line 24 to page 55, line 19 and Fig. 20. The spelling rules are employed in identifying word fragments. See the specification at page 47, line 29 to page 48, line 14; page 51, lines 3-7

Independent claim 46 is directed to a computer-implemented speech recognition system that uses an expanded effective active vocabulary. The system includes a storage device configured to store an active vocabulary that includes multiple entries corresponding to words, commands, and word fragments. The system includes a processor configured to perform the method of claim 1, and is described in the specification by the passages noted above with respect to claim 1.

Independent claim 51 is directed to computer software, residing on a computer readable medium, for a speech recognition system that uses an expanded effective active vocabulary to recognize words, and commands. The computer software includes instructions for causing a

computer to perform the method of claim 1 and is described in the specification by the passages noted above with respect to claim 1.

(6) Grounds of rejection to be reviewed on appeal

Claims 1, 2, 11-21, 30-33, 36, 37, and 46-55 stand rejected as being unpatentable over U.S. Patent No. 6,212,498 (Sherwood) in view of U.S. Patent No. 5,765,132 (Roberts) and U.S. Patent No. 6,092,044 (Baker). Claims 3-9, 22-29, 34, and 38-45 stand rejected as being unpatentable over Sherwood in view of Roberts, Baker, and U.S. Patent No. 5,835,888 (Kanevsky). Claim 35 stands rejected as being obvious over Sherwood in view of Roberts.

(7) Argument

The subject matter of claims 1, 2, 11-21, 30-33, 36, 37, and 46-55 would not have been obvious over Sherwood in view of Roberts and Baker.

Claims 1, 46, and 51:

Appellant requests reversal of the rejection of claims 1, 46, and 51 because neither Sherwood, Roberts, Baker, nor any proper combination of the three describes or suggests (1) combining a particular word fragment with one or more adjacent word fragments using an associated spelling rule to form a proposed word having a spelling that differs from a spelling that would result from merely concatenating the word fragment with one or more adjacent word fragments or words and (2) when the word fragment may not be combined with adjacent word fragments or words to form a proposed word, discarding the recognition candidate that includes the word fragment.

As the Examiner agrees, neither Sherwood nor Roberts describes or suggests spelling rules. Moreover, Baker fails to cure the deficiencies of Sherwood to describe or suggest such spelling rules. Baker's system permits a user to add a new word to the dictation vocabulary by typing in the word and then uttering the word. See Baker at col. 15, lines 56-60. Baker's system includes a list that matches each string of letters with possible phonemes. See Baker at col. 15, lines 60-63. Baker's system then creates a set of words, with each word representing a possible phonetic spelling based on the phoneme/letter string list. See Baker at col. 15, lines 63-65 and col. 16, lines 1-5.

However, Baker's phoneme/letter string list does not include or relate to spelling rules. The only rules present in Baker's phoneme/letter string list are phoneme rules that associate each letter string with a particular list of phonemes. See Baker at col. 15, lines 56-65. Neither the entries in the list nor the list itself includes a spelling rule. Furthermore, not only is Baker silent on spelling rules, Baker's system would receive no benefit from the use of spelling rules since the spelling of the word is input into Baker's system by the user to generate the letter strings.

Moreover, even if one were to modify Sherwood to incorporate the phoneme/letter string list of Baker, Sherwood would still fail to form a proposed word having a spelling that differs from a spelling that would result from merely concatenating a particular word fragment with the one or more adjacent word fragments or words. Baker is silent as to the concept of producing a word having a spelling that differs from the spelling that would result from concatenating a word fragment with one or more adjacent word fragments or words. One reason for this silence is that Baker does not address word fragments. Baker relates to a method for adding a word to a vocabulary based on a spelling of the word and not based on fragments of the word. See Baker at abstract. Another reason for this silence is that Baker receives the exact spelling from the user and there is no opportunity to concatenate a word fragment with one or more adjacent word fragments or words. See Baker at col. 15, lines 56-65. For this reason, Baker cannot remedy the failure of Sherwood and Roberts to describe or suggest combining a particular word fragment with one or more adjacent word fragments using an associated spelling rule to form a proposed word having a spelling that differs from a spelling that would result from merely concatenating the word fragment with one or more adjacent word fragments or words, as recited in each of independent claims 1, 46, and 51.

Sherwood, Roberts, and Baker also are silent with respect to discarding a recognition candidate if a word fragment may not be combined with one or more adjacent word fragments or words to form a proposed word. Furthermore, the Examiner does not point to any features of Sherwood, Roberts, or Baker that might show such discarding of a recognition candidate.

For at least the reasons provided above, appellant requests reversal of the rejection of claims 1, 46, and 51.

Claims 2, 11-21, 30-33, 47-50, and 52-55:

Claims 2, 11-21, 30-33, 47-50, and 52-55 depend from claims 1, 46, or 51, which stand rejected as being unpatentable over Sherwood in view of Roberts and Baker. Appellant requests reversal of the rejection of claims 2, 11-21, 30-33, 36, 37, 47-50, and 52-55 because these claims are allowable for at least the reasons that claims 1, 46, and 51 are allowable and for containing allowable subject matter in their own right.

For example, claim 2 recites that the expanded effective vocabulary includes words from the backup dictionary and words from the active vocabulary. The words from the backup dictionary are formed from a combination of word fragments and either words or word fragments from an active vocabulary that includes words and word fragments. Neither Sherwood, Roberts, nor Baker describes or suggests an expanded effective vocabulary that includes words from a backup dictionary that are formed from a combination of words and word fragments or word fragments and word fragments from an active vocabulary. Moreover, neither Sherwood, Roberts, nor Baker combines word fragments with other word fragments or with words. Realizing that Sherwood and Baker are silent as to the combination of word fragments, the Examiner points to Roberts as somehow showing word fragments. The Examiner explains [emphasis added]:

> Roberts teaches generating recognition candidates for a received user utterance, recognizing fragments of an utterance, modifying the recognition candidate and searching the vocabulary for the modified recognition candidate before adding the word to the vocabulary. Roberts teaches the speech recognition system builds speech models for new words without requiring the user to discretely speak the new word such that addition of a new word to the system vocabulary appears as a simple correction of a mis-recognized word.

See the Final Action at Page 4. However, while Roberts may recognize fragments of an utterance, Roberts fails to suggest a fragment of a word. Roberts only deals in complete words and never suggests combining a word fragment with adjacent word fragments or words. Roberts explains in the abstract [emphases added]

> [W]ords are added to a speech recognition system vocabulary during user dictation by (a) extracting, from a multi-word user utterance, speech frames that correspond to each one of the one or more new words; and (b) building speech models for the one or more new words using the extracted speech frames.

As another example, claim 33 recites that the determination of whether a word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in the backup dictionary of the speech recognition system includes searching the backup dictionary for the proposed word based on a pronunciation of the proposed word. Neither Sherwood, Roberts, nor Baker describes or suggests searching a backup dictionary for a proposed word formed from a combination of a word fragment with one or more adjacent word fragments or words. As discussed above, Roberts never determines whether a word fragment may be combined with an adjacent word fragment to form a proposed word included in a backup dictionary.

For at least these reasons, appellant requests reversal of the rejection of claims 2, 11-21, 30-33, 47-50, and 52-55.

Claim 36:

Claim 36 depends from claim 35, which stands rejected as being obvious over Sherwood in view of Roberts. Appellant addresses the rejection of claim 36 after addressing the rejection of claim 35 below.

Claim 37:

Appellant requests reversal of the rejection of independent claim 37 because neither Sherwood, Roberts, Baker, nor any proper combination of the three describes or suggests comparing words of an active vocabulary to similar words of a backup dictionary to identify spelling rules that may be used to convert the words of the active vocabulary to words of the backup dictionary, or employing the spelling rules in identifying word fragments.

As previously discussed, Sherwood and Roberts are silent as to identifying spelling rules and using spelling rules to identify word fragments. Additionally, as discussed above, Baker does not identify spelling rules. Instead, Baker's system relies on phoneme rules. See Baker at col. 15, lines 56-65.

Even assuming for the sake of argument that Baker describes the use of spelling rules, Baker nowhere describes or suggests identifying spelling rules that may be used to convert words

of an active vocabulary to words of a backup dictionary, and Baker also nowhere describes or suggests employing the spelling rules to identify word fragments, as recited in claim 37. In Baker's system, a user spells and speaks a word to be added to a vocabulary and the spoken word is recognized against the constraint grammar. Baker's system does not identify spelling rules to convert words from an active vocabulary to words of a backup dictionary, and does not employ such spelling rules to identify fragments of words. Accordingly, appellant requests reversal of the rejection of claim 37.

The subject matter of claims 3-9, 22-29, 34, and 38-45 would not have been obvious over Sherwood in view of Roberts, Baker, and Kanevsky.

Claims 3-9 and 22-29:

Claims 3-9 and 22-29 depend from independent claim 1, which stands rejected as being obvious over Sherwood in view of Roberts and Baker. Appellant requests reversal of the rejection of claims 3-9 and 22-29 because neither Sherwood, Roberts, Baker, nor any proper combination of the three describes or suggests (1) combining a particular word fragment with one or more adjacent word fragments using an associated spelling rule to form a proposed word having a spelling that differs from a spelling that would result from merely concatenating the word fragment with one or more adjacent word fragments or words and (2) when the word fragment may not be combined with adjacent word fragments or words to form a proposed word, discarding the recognition candidate that includes the word fragment, and because Kanevsky fails to remedy the deficiencies of Sherwood, Roberts, and Baker.

Kanevsky relates to a system for generating a language model that includes a vocabulary and vocabulary components such as stems, endings, and prefixes, which are formed from the vocabulary. See Kanevsky at col. 4, lines 14-20. The vocabulary components are produced by a splitter that splits the words in the vocabulary. See Kanevsky at col. 4, lines 20-23. Kanevsky generates the language model by calculating a probability of a word as a "weighted sum of output probabilities of several" language models, which are built on vocabulary components and combinations of vocabulary components. See Kanevsky at col. 3, lines 6-17.

However, while Kanevsky discusses vocabulary components, Kanevsky fails to show combining vocabulary components to form a proposed word using a spelling rule that produces a

spelling of the proposed word that differs from a spelling that would result from merely concatenating the word component with other word components. Rather, as Kanevsky explains, the vocabulary components are merely concatenated: "stems are connected consequently ... and matched ... to see which concatenations of stems produce existing words in the vocabulary." See Kanevsky at col. 7, lines 16-21 and Fig. 7.

Furthermore, Kanevsky also is silent with respect to discarding a recognition candidate if a word fragment may not be combined with one or more adjacent word fragments or words to form a proposed word. Kanevsky merely points out that "allowable sequences of words" are produced by the connection of the stems. See Kanevsky at col. 7, lines 20-22.

Accordingly, claim 1 is allowable over any proper combination of Sherwood, Roberts, Baker, and Kanevsky. Claims 3-9 and 22-29 depend from claim 1 and are allowable for at least the reasons that claim 1 is allowable. Accordingly, appellant requests reversal of the rejection of claims 3-9 and 22-29.

Claim 34:

As discussed above with respect to claims 3-9 and 22-29, neither Sherwood, Roberts, Baker, Kanevsky nor any combination of the four describes or suggests combining a particular word fragment with one or more adjacent word fragments using an associated spelling rule in forming the proposed word and a spelling of a proposed word differing from a spelling that would result from merely concatenating the word fragment with the one or more adjacent word fragments or words. For this reason, the combination of Sherwood, Roberts, Baker, and Kanevsky also fails to describe or suggest formation of a new word using a spelling rule associated with an affix that causes the spelling of the new word to differ from a spelling that would result from merely concatenating the affix with the one or more adjacent words, root, or other affixes that form the new word, as recited in claim 34. Accordingly, appellant requests reversal of the rejection of claim 34.

Claims 38-45:

Claims 38-45 depend from independent claim 37, which stands rejected as being obvious over Sherwood in view of Roberts and Baker. As discussed above, neither Sherwood, Roberts,

Baker, nor any proper combination of the three describes or suggests comparing words of an active vocabulary to similar words of a backup dictionary to identify spelling rules that may be used to convert the words of the active vocabulary to words of the backup dictionary and employing the spelling rules in identifying word fragments, as recited in claim 37. Moreover, Kanevsky fails to remedy this deficiency of Sherwood, Roberts, and Baker.

As discussed above, Kanevsky fails to show combining vocabulary components to form a proposed word using a spelling rule. Furthermore, Kanevsky also fails to show comparing words of an active vocabulary to similar words of a backup dictionary to identify spelling rules. Kanevsky connects strings of stems and endings and the connected strings are matched with the vocabulary to determine "which concatenations of stems produce existing words in the vocabulary." See Kanevsky at col. 7, lines 16-22 and Fig. 7. However, Kanevsky's connected strings are not in an active vocabulary and Kanevsky never compares the connected strings to words in the vocabulary to identify spelling rules. Rather, Kanevsky merely matches the stems with the vocabulary to produce "allowable sequences of words." See Kanevsky at col. 7, lines 16-22 and Fig. 7. Lastly, because Kanevsky fails to identify spelling rules, Kanevsky also fails to employ spelling rules, as also required by claim 37. For this reason, claim 37 is allowable over any possible combination of Sherwood, Roberts, Baker, and Kanevsky.

Claims 38-45 depend from claim 37 and are allowable for at least the reasons that claim 37 is allowable and for containing allowable subject matter in their own right. For example, claim 38 recites that employing spelling rules in identifying word fragments includes grouping spelling rules together to form possible affixes and analyzing words of the backup dictionary using the affixes to identify roots that may be combined with the affixes to produce words of the backup dictionary. Because none of the cited art identifies spelling rules, none of the cited art groups spelling rules together to form possible affixes.

As another example, claim 41 recites that the method also includes storing a set of spelling rules in association with an affix in the active vocabulary. Because none of the cited art identifies spelling rules, none of the cited art stores a set of spelling rules.

For these reasons, appellant requests reversal of the rejection of claims 38-45.

The subject matter of claim 35 would not have been obvious over Sherwood in view of Roberts.

Claim 35:

Appellant requests reversal of the rejection of claim 35 because any theoretical combination of Sherwood and Roberts would still fail to describe or suggest comparing a word of an active vocabulary to a similar word of a backup dictionary to identify a word fragment that may be used to convert the word of the active vocabulary to the word of the backup dictionary, and would also fail to describe or suggest generating the acoustic model of the word fragment using a portion of an acoustic model of the word of the backup dictionary that is not included in an acoustic model of the word of the active vocabulary, as recited in claim 35.

As the Examiner agrees, Sherwood does not relate to word fragments. Thus, Sherwood cannot describe or suggest identifying a word fragment and generating an acoustic model of the word fragment.

Moreover, Roberts fails to cure the deficiencies of Sherwood. Roberts' speech recognition system does not identify a word fragment that may be used to convert a word of an active vocabulary to a word of a backup dictionary, and does not generate an acoustic model of the word fragment using a portion of an acoustic model of the word of the backup dictionary that is not included in the acoustic model of the word of the active vocabulary. While, as the Examiner notes, Roberts teaches the identification of fragments or segments of a phrase "This is a black cow," Roberts does not show identification of word fragments. Moreover, Roberts identifies fragments of the phrase by analyzing the speech frames of an utterance, and does not identify phrase fragments by comparing the speech frames of a word from an active vocabulary to the speech frames of a word from a backup dictionary. Roberts explains "the speech recognition system recognizes (step 52) the best candidate as a combination of words corresponding to speech models that most closely match the speech frames" and "the speech recognition system ... attempts to isolate the speech frames corresponding to the new word so that the speech frames may be extracted and used to build a speech models for the new word." See Roberts at col. 5, lines 11-25. Thus, as noted previously, Roberts' speech recognition system merely determines that a corrected word is not in the vocabulary and then isolates the speech

frames for the new word from the speech frames for the utterance in order to build a speech model for the new word. Accordingly, the rejection should be reversed.

Claim 36:

Claim 36, which stands rejected as being obvious over Sherwood in view of Roberts and Baker, depends from claim 35, which stands rejected as being obvious over Sherwood in view of Roberts. As discussed above, neither Sherwood, Roberts, nor any proper combination of the two describes or suggests comparing a word of an active vocabulary to a similar word of a backup dictionary to identify a word fragment that may be used to convert the word of the active vocabulary to the word of the backup dictionary, and would also fail to describe or suggest generating the acoustic model of the word fragment using a portion of an acoustic model of the word of the backup dictionary that is not included in an acoustic model of the word of the active vocabulary, as recited in claim 35.

Baker fails to cure the deficiencies of Sherwood and Roberts. Baker never compares a word of an active vocabulary to a similar word of a backup dictionary to identify a word fragment. In Baker, a collection of phonetic pronunciations is created based on a user input of a spelling and an utterance of a word. See Baker at col. 1, lines 49-53. Baker's system then finds the pronunciation from the collection that best matches the utterance. See Baker at col. 1, lines 53-54. The word is added to the "vocabulary using the spelling and the best-matching pronunciation." See Baker at col. 1, lines 54-57. However, Baker never compares the word to a similar word of a backup dictionary to identify a word fragment. See Baker at Fig. 13. Accordingly, any possible combination of Sherwood, Roberts, and Baker would still fail to describe or suggest comparing a word of an active vocabulary to a similar word of a backup dictionary to identify a word fragment, and would also inherently fail to describe or suggest generating an acoustic model of the word fragment, as recited in claim 35. Thus, claim 35 is allowable over Sherwood, Roberts, and Baker.

Claim 36 is allowable for at least the reasons that claim 35 is allowable and for containing allowable subject matter in its own right. Claim 36 recites that the comparing of a word of the active vocabulary to a similar word of the backup dictionary includes comparing spellings of the two words. Neither Sherwood, Roberts, Baker, nor any proper combination of

the three describes or suggests comparing a spelling of a word of an active vocabulary to a spelling of a word of a backup dictionary. Sherwood never mentions a spelling of a word. In Roberts, while a user changes the spelling of an incorrectly recognized word (for example, changing the spelling of "how" to "cow") the system never compares the spelling of the word to a spelling of a word in a backup dictionary and the word itself is not in an active vocabulary since it is a new word. See Roberts at col. 5, lines 11-21. Similarly, in Baker, although the user inputs a spelling of a word, Baker's system never compares that spelling to a spelling of a word in a backup dictionary and the word itself is not in an active vocabulary since the word is a new word. See Baker at col. 15, line 56 to col. 16, line 5.

For these reasons, appellant requests the withdrawal of the rejection of claim 36.

The Examiner failed to address appellant's arguments in the Advisory Action dated October 22, 2004.

Many of the arguments presented above were raised in reply to the final action. However, in the advisory action mailed four months after the reply, the Examiner cites In re Keller, 642 F.2d 413, 208 USPQ 871 (CCPA 1981) and In re Merck & Co., 800 F.2d 1091, 231 USPQ 375 (Fed. Cir. 1986) and merely states "one cannot show nonobviousness by attacking references individually where the rejections are based on combinations of references." Apparently, the Examiner never considered appellant's arguments (1) because the Examiner failed to reach the merits of appellant's arguments in the advisory action and (2) because the Examiner failed to address appellant's arguments that any possible combination of the references still failed to describe or suggest features of each of the claims. For this reason, the Examiner failed to establish a prima facie case of obviousness, which requires, among other things, that the references, when combined, teach or suggest all of the claim limitations.

Conclusion

For the foregoing reasons, the rejections should be reversed.

This appeal brief is being filed within seven months from April 29, 2004, the filing date of the Notice of Appeal. Enclosed is a check to cover the fee for this brief and a check to cover the extension of time fee. Please apply any other charges or credits to Deposit Account No. 06-1050.

Respectfully submitted,

Date: November 10, 2004

Diana DiBerardino

Diana DiBerardino
Reg. No. 45,653

Fish & Richardson P.C.
1425 K Street, N.W.
11th Floor
Washington, DC 20005-3500
Telephone: (202) 783-5070
Facsimile: (202) 783-2331

Appendix of Claims

1. (Previously presented) A method of expanding an effective active vocabulary of a speech recognition system, the method comprising:

using a speech recognizer to perform speech recognition on a user utterance to produce one or more recognition candidates, the speech recognition comprising comparing digital values representative of the user utterance to a set of acoustic models representative of an active vocabulary of the system, the set of acoustic models including models of words and models of word fragments,

receiving the recognition candidates from the speech recognizer, and

when a received recognition candidate includes a word fragment:

determining whether the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system, wherein forming the proposed word includes using a spelling rule associated with the word fragment that causes the spelling of the proposed word to differ from a spelling that would result from merely concatenating the particular word fragment with the one or more adjacent word fragments or words;

if the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system, modifying the recognition candidate to substitute the proposed word for the word fragment and the one or more adjacent word fragments or words used to form the proposed word; and

if the word fragment may not be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system, discarding the recognition candidate.

2. (Original) The method of claim 1, wherein the expanded effective vocabulary comprises words from the backup dictionary that are formed from a combination of words and word fragments or word fragments and word fragments from an active vocabulary that includes words and word fragments, and words from the active vocabulary.

3. (Original) The method of claim 1, wherein word fragments comprise suffixes, prefixes, and roots that are not words.

4. (Original) The method of claim 3, wherein:

one or more spelling rules are associated with each prefix and each suffix,

determining whether the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word comprises using a prefix or suffix as the particular word fragment and using an associated spelling rule in forming the proposed word, and

as a result of using the associated spelling rule, a spelling of the proposed word differs from a spelling that would result from merely concatenating the particular word fragment with the one or more adjacent word fragments or words.

5. (Original) The method of claim 4, wherein determining whether the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word comprises:

retrieving from the received recognition candidate a sequence that includes the particular word fragment and adjacent word fragments or words; and

determining if the sequence is a valid sequence.

6. (Original) The method of claim 5, wherein a valid sequence includes only one or more allowed adjacent combinations of word fragments and words.

7. (Original) The method of claim 6, wherein allowed adjacent combinations comprise one or more prefixes, followed by a root or a word, followed by one or more suffixes; a root or a word followed by one or more suffixes; and one or more prefixes followed by a root or a word.

8. (Original) The method of claim 6, wherein allowed adjacent combinations comprise one or more prefixes, followed by one or more roots or words, followed by one or more suffixes; one or more roots or words followed by one or more suffixes; and one or more prefixes followed by one or more roots or words.

9. (Original) The method of claim 4, further comprising combining the particular word fragment with the one or more adjacent word fragments or words to form a second

proposed word that differs from the first proposed word by using a second associated spelling rule in forming the proposed word.

10. (Canceled)

11. (Original) The method of claim 1, wherein determining whether the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system comprises searching the backup dictionary for the proposed word.

12. (Original) The method of claim 1, wherein modifying the recognition candidate comprises:

forming a prospective recognition candidate by modifying the recognition candidate to substitute the proposed word for the word fragment and the one or more adjacent word fragments or words used to form the proposed word; and

if the prospective recognition candidate includes an additional word fragment:

further processing the prospective recognition candidate to generate an additional word using the additional word fragment and one or more adjacent words or word fragments, and

forming a final recognition candidate by replacing the additional word fragment and the one or more adjacent words with the additional word.

13. (Original) The method of claim 1, wherein a score is associated with the received recognition candidate, the method further comprising producing a score associated with the modified recognition candidate by rescoring the modified recognition candidate.

14. (Original) The method of claim 13, wherein:

the score associated with the received recognition candidate includes an acoustic component and a language model component, and

rescoring the modified recognition candidate comprises generating a language model score for the modified recognition candidate.

15. (Original) The method of claim 14, wherein producing the score associated with the modified recognition candidate comprises combining the acoustic component of the score for the received recognition candidate with the language model score generated for the modified recognition candidate.

16. (Original) The method of claim 14, wherein rescoring the modified recognition candidate comprises generating an acoustic model score for the modified recognition candidate and producing the score associated with the modified recognition candidate comprises combining the acoustic model score generated for the modified recognition candidate with the language model score generated for the modified recognition candidate.

17. (Original) The method of claim 13, wherein:

the score associated with the received recognition candidate includes an acoustic component and a language model component; and

rescoring the modified recognition candidate comprises generating an acoustic score for the modified recognition candidate.

18. (Original) The method of claim 17, wherein producing the score associated with the modified recognition candidate comprises combining the language model component of the score for the received recognition candidate with the acoustic score generated for the modified recognition candidate.

19. (Original) The method of claim 1, further comprising generating the acoustic model of a word fragment by:

comparing a word of the active vocabulary to a similar word of a backup dictionary to identify a word fragment that may be used to convert the word of the active vocabulary to the word of the backup dictionary, and

generating the acoustic model of the word fragment using a portion of an acoustic model of the word of the backup dictionary that is not included in an acoustic model of the word of the active vocabulary.

20. (Original) The method of claim 19, wherein comparing a word of the active vocabulary to a similar word of a backup dictionary comprises comparing spellings of the two words.

21. (Original) The method of claim 1, further comprising generating the acoustic models of word fragments by:

comparing words of the active vocabulary to similar words of a backup dictionary to identify spelling rules that may be used to convert the words of the active vocabulary to words of the backup dictionary; and

employing the spelling rules in identifying word fragments.

22. (Original) The method of claim 21, wherein employing the spelling rules in identifying the word fragments comprises:

grouping spelling rules together to form possible affixes, the affixes including prefixes and suffixes; and

analyzing words of the backup dictionary using the affixes to identify roots that may be combined with the affixes to produce words of the backup dictionary.

23. (Original) The method of claim 22, further comprising generating acoustic models for the roots using portions of acoustic models of the words of the backup dictionary that are not included in acoustic models of the affixes.

24. (Original) The method of claim 22, further comprising adding affixes and roots to the active vocabulary as word fragments.

25. (Original) The method of claim 24, further comprising storing a set of spelling rules in association with an affix in the active vocabulary.

26. (Original) The method of claim 24, further comprising creating a language model associated with the active vocabulary.

27. (Original) The method of claim 26, wherein creating the language model comprises:

retrieving a training collection of text, the training collection of text comprising words from the backup dictionary and words from the active vocabulary;

modifying the training collection of text by replacing any splittable backup dictionary words with their corresponding words and word fragments; and

generating language model scores for words and word fragments of the active vocabulary using the modified collection of text.

28. (Original) The method of claim 26, wherein creating the language model comprises creating an N-gram language model.

29. (Original) The method of claim 28, wherein creating the N-gram language model comprises:

retrieving a training collection of text, the training collection of text comprising words from the backup dictionary and words from the active vocabulary;

determining a frequency of each N-gram word sequence that appears in the training collection of text;

modifying the N-gram word sequences by replacing any splittable backup dictionary words with their corresponding words and word fragments;

based on the N-gram word sequence frequencies, determining a frequency of each modified N-gram sequence that includes words, word fragments, or words and word fragments; and

based on the N-gram word and word fragment sequence frequencies, generating an N-gram language model for the words and word fragments of the active vocabulary.

30. (Original) The method of claim 1, wherein word fragments comprise syllables, and each syllable includes a unit of spoken language.

31. (Original) The method of claim 30, wherein the unit of spoken language comprises a single uninterrupted sound formed by a vowel or diphthong alone, of a syllabic consonant alone, or of either with one or more consonants.

32. (Original) The method of claim 30, wherein a syllable comprises a vowel and one or more consonants clustered around the vowel.

33. (Original) The method of claim 1, wherein determining whether the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in the backup dictionary of the speech recognition system comprises searching the backup dictionary for the proposed word based on a pronunciation of the proposed word.

34. (Previously presented) A method of recognizing speech, the method comprising:

using a speech recognizer to perform speech recognition on a user utterance to produce a set of one or more recognition candidates, the speech recognition comprising comparing digital values representative of the user utterance to a set of acoustic models representative of an active vocabulary of the system, the set of acoustic models including models of words, models of roots that are not words, and models of affixes that are not words, the affixes including prefixes and suffixes,

receiving the recognition candidates from the speech recognizer, and

when a received recognition candidate includes an affix:

combining the affix with one or more adjacent words, roots, or other affixes to form a new word, wherein forming the new word includes using a spelling rule associated with the affix that causes the spelling of the new word to differ from a spelling that would result from merely concatenating the affix with the one or more adjacent words, roots, or other affixes; and

modifying the recognition candidate to substitute the new word for the affix and the one or more adjacent words, roots, or other affixes used to form the new word.

35. (Original) A method of generating an acoustic model of a word fragment, the method comprising:

comparing a word of an active vocabulary to a similar word of a backup dictionary to identify a word fragment that may be used to convert the word of the active vocabulary to the word of the backup dictionary, and

generating the acoustic model of the word fragment using a portion of an acoustic model of the word of the backup dictionary that is not included in an acoustic model of the word of the active vocabulary.

36. (Original) The method of claim 35, wherein comparing a word of the active vocabulary to a similar word of a backup dictionary comprises comparing spellings of the two words.

37. (Original) A method of generating acoustic models of word fragments, the method comprising:

comparing words of an active vocabulary to similar words of a backup dictionary to identify spelling rules that may be used to convert the words of the active vocabulary to words of the backup dictionary; and

employing the spelling rules in identifying word fragments.

38. (Original) The method of claim 37, wherein employing the spelling rules in identifying word fragments comprises:

grouping spelling rules together to form possible affixes, the affixes including prefixes and suffixes; and

analyzing words of the backup dictionary using the affixes to identify roots that may be combined with the affixes to produce words of the backup dictionary.

39. (Original) The method of claim 38, further comprising generating acoustic models for the roots using portions of acoustic models of the words of the backup dictionary that are not included in acoustic models of the affixes.

40. (Original) The method of claim 38, further comprising adding affixes and roots to the active vocabulary as word fragments.

41. (Original) The method of claim 40, further comprising storing a set of spelling rules in association with an affix in the active vocabulary.

42. (Original) The method of claim 40, further comprising creating a language model associated with the active vocabulary.

43. (Original) The method of claim 42, wherein creating the language model comprises:

retrieving a training collection of text, the training collection of text comprising words from the backup dictionary and words from the active vocabulary;

modifying the training collection of text by replacing any splittable backup dictionary words with their corresponding words and word fragments; and

generating language model scores for words and word fragments of the active vocabulary using the modified collection of text.

44. (Original) The method of claim 42, wherein creating the language model comprises creating an N-gram language model.

45. (Original) The method of claim 44, wherein creating the N-gram language model comprises:

retrieving a training collection of text, the training collection of text comprising words from the backup dictionary and words from the active vocabulary;

determining a frequency of each N-gram word sequence that appears in the training collection of text;

modifying the N-gram word sequences by replacing any splittable backup dictionary words with their corresponding words and word fragments;

based on the N-gram word sequence frequencies, determining a frequency of each modified N-gram sequence that includes words, word fragments, or words and word fragments; and

based on the N-gram word and word fragment sequence frequencies, generating an N-gram language model for the words and word fragments of the active vocabulary.

46. (Previously presented) A computer-implemented speech recognition system that uses an expanded effective active vocabulary, the system comprising:

a storage device configured to store an active vocabulary that includes multiple entries corresponding to words, commands, and word fragments; and

a processor configured to:

receive data representing a user utterance,

produce one or more recognition candidates, by comparing digital values representative of the user utterance to a set of acoustic models representative of the active vocabulary of the system, the set of acoustic models including models of words and models of word fragments,

when a produced recognition candidate includes a word fragment:

determine whether the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system, wherein forming the proposed word includes using a spelling rule associated with the word fragment that causes the spelling of the proposed word to differ from a

spelling that would result from merely concatenating the particular word fragment with the one or more adjacent word fragments or words;

if the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary, modify the recognition candidate to substitute the proposed word for the word fragment and the one or more adjacent word fragments or words used to form the proposed word; and

if the word fragment may not be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system, discard the recognition candidate.

47. (Original) The system of claim 46, wherein the expanded effective active vocabulary comprises words from the backup dictionary that are formed from a combination of words and word fragments or word fragments and word fragments from an active vocabulary that includes words and word fragments, and words from the active vocabulary.

48. (Original) The system of claim 46, wherein the processor determines whether the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in the backup dictionary by searching the backup dictionary for the proposed word.

49. (Original) The system of claim 46, wherein the processor modifies the recognition candidate by:

forming a prospective recognition candidate by modifying the recognition candidate to substitute the proposed word for the word fragment and the one or more adjacent word fragments or words used to form the proposed word; and

if the prospective recognition candidate includes an additional word fragment:

further processing the prospective recognition candidate to generate an additional word using the additional word fragment and one or more adjacent words or word fragments, and

forming a final recognition candidate by replacing the additional word fragment and the one or more adjacent words with the additional word.

50. (Original) The system of claim 46, wherein a score is associated with the received recognition candidate, and the processor is further configured to produce a score associated with the modified recognition candidate by rescoring the modified recognition candidate.

51. (Previously presented) Computer software, residing on a computer readable medium, for a speech recognition system that uses an expanded effective active vocabulary to recognize words, and commands, the computer software comprising instructions for causing a computer to perform the following operations:

receive data representing a user utterance,

produce one or more recognition candidates, by comparing digital values representative of the user utterance to a set of acoustic models representative of an active vocabulary of the system, the set of acoustic models including models of words and models of word fragments,

when a produced recognition candidate includes a word fragment:

determine whether the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system, wherein forming the proposed word includes using a spelling rule associated with the word fragment that causes the spelling of the proposed word to differ from a spelling that would result from merely concatenating the particular word fragment with the one or more adjacent word fragments or words;

if the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary, modify the recognition candidate to substitute the proposed word for the word fragment and the one or more adjacent word fragments or words used to form the proposed word; and

if the word fragment may not be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system, discard the recognition candidate.

52. (Original) The computer software of claim 51, wherein the expanded effective active vocabulary comprises words from the backup dictionary that are formed from a combination of words and word fragments or word fragments and word fragments from an active vocabulary that includes words and word fragments, and words from the active vocabulary.

53. (Original) The computer software of claim 51, wherein determining whether the word fragment may be combined with one or more adjacent word fragments or words to form a

proposed word included in the backup dictionary comprises searching the backup dictionary for the proposed word.

54. (Original) The computer software of claim 51, wherein modifying the recognition candidate comprises:

forming a prospective recognition candidate by modifying the recognition candidate to substitute the proposed word for the word fragment and the one or more adjacent word fragments or words used to form the proposed word; and

if the prospective recognition candidate includes an additional word fragment:

further processing the prospective recognition candidate to generate an additional word using the additional word fragment and one or more adjacent words or word fragments, and

forming a final recognition candidate by replacing the additional word fragment and the one or more adjacent words with the additional word.

55. (Original) The computer software of claim 51, wherein a score is associated with the received recognition candidate, the computer software comprising instructions for causing the computer to produce a score associated with the modified recognition candidate by rescoring the modified recognition candidate.